

ENTERED



1600

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/993,192

DATE: 07/30/2003
 TIME: 15:47

RECEIVED

Aug 1 2003
 TECH CENTER 1600/2900

Input Set : N:\Crf3\RULE60\09993192.raw.txt
 Output Set: N:\CRF4\07302003\I993192.raw

of

```

1 <110> APPLICANT: Korea Research Institute of Bioscience and Biotechnology
2      Dong Kook Pharmaceutical Co.
3      Rhee, Sangki
4      Choi, Euisung
5      Kang, Hyunah
6      Sohn, Junghoon
7      Bae, Junghoon
8      Kim, Moowoong
9      Agaphonov, Michasel
10 <120> TITLE OF INVENTION: Hansenula polymorpha mutants and process for the preparation
of
11      recombinant proteins using the same
12 <130> FILE REFERENCE: 4220-116 US
13 <140> CURRENT APPLICATION NUMBER: US/09/993,192
14 <141> CURRENT FILING DATE: 2001-11-14
15 <150> PRIOR APPLICATION NUMBER: US/09/674,617A
16 <151> PRIOR FILING DATE: 2001-01-03
17 <150> PRIOR APPLICATION NUMBER: PCT/KR00/00173
18 <151> PRIOR FILING DATE: 2000-03-04
19 <160> NUMBER OF SEQ ID NOS: 9
20 <170> SOFTWARE: PatentIn version 3.0
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 2218
24 <212> TYPE: DNA
25 <213> ORGANISM: Hansenula polymorpha DL1
26 <300> PUBLICATION INFORMATION:
27 <308> DATABASE ACCESSION NO: U67174
28 <309> DATABASE ENTRY DATE: 1996-08-17
29 <313> RELEVANT RESIDUES: (1)..(2218)
30 <400> SEQUENCE: 1
31      ctcgagtatg attactttgt cttaatTTTT atttgggtga attttttact accatttttaa      60
32      tgacgtgggtg tacgggtgtg aatacataaa cacctcagta gccatagcag ctaccatgga      120
33      tgatagggca ctttttgaaa aggaatctga aaacaagaag aagcacggca tacgggtggtc      180
34      aacagctcaa gaaaggccaa tgcgataaat ctcattttta tacgaacacc caccgaagcc      240
35      cttatcggtc ttttgacgca gcagcttaaat tatctgaggc tggttatcaa tttttgcctt      300
36      ccagataaaa tatattttcc ctattttatat tccatccggt tatcctaggc aatcgtccaa      360
37      aaaaacaaga attctgttat atttcaattc ccattatgaa gctctcgatg tctattctgg      420
38      ccctagtggc cactagcatt gggtcggcgc aagccgctgc catcaaaaaa gataccgcgg      480
39      ggcaacaccc cctcggaatg aactccaact ttggcgatgt tttcgtggag aaaggcaagt      540
40      ctttgctcga ccgtgtttct gaggttgta gcgagtcgct caagaatatt tcccagaagt      600
41      tgagggacat ttggaatgaa atggaggcca agttcccaga taccctcaga aacatgaagc      660
42      tcaagtcaga gccggcagtc aaaattacaa agaaacctgc cgatttttgg gacttcaatg      720
43      ttctcaatga gaagttctcc aactacaagc tgagggttaa gaagaccgac ccgggagcat      780
44      tgggactgga ccacacaaga cagtactctg gatacttgga tgtggaggac gaagacaagc      840

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/993,192

DATE: 07/30/2003
TIME: 10:45:47

Input Set : N:\Crf3\RULE60\09993192.raw.txt
Output Set: N:\CRF4\07302003\I993192.raw

```

45 atttcttota ttggatgttt gagtccagaa atgaccgcgt caacgacccgt gtgattctgt 900
46 ggctcaacgg ttgtccagga tgcctcttct tgactggaat gctttttgag ctccggtctgt 960
47 ctctctatcg tccagatctc aagccaatca acaaccata tctgtggaat tccaatgcc 1020
48 ctgtgatttt ccttgaccag cctgtcaatg ttggattctc gtactcttcc aagctctgtt 1080
49 ctaacacggt cgcagctggt aaagacgtct atgctttctt ggagttgttc taccagcaat 1140
50 tcccacactt gctgaagaac gacttccaca tcgcgcgcga gtcgtacggt ggtcattaca 1200
51 tcccagtggt tgctcccgag attctcacc atgctgacag atctttcaac ctcaactcgg 1260
52 tgttgattgg taacggtttg accgacccac ttaaccagta cccattctac gagagaatgg 1320
53 catgctctac tgatgggtgc tatgagcaac cctggacgag tctgagtgcg aaggaaatgtt 1380
54 ttggagacct gccatagatt ttgtcattga ttgaatcatg ctacagctcg cagctctgtg 1440
55 tctcatgtgt cccggcctcc atctactgca acaacgcaca acttggacca ttccaaaaga 1500
56 ccggcagaaa cgtctacgac gttagaaga ttgtcgcagg aactctgtgc tacaagaca 1560
57 ttgaatacat tgaccaatat ttgaaccagg actttgtcaa ggaaagggtt ggcgctgagg 1620
58 ttgacactta cgagtcgtgt aatttcgacg tgaacagaaa ctctctgttt gctgggtgatt 1680
59 ggatgaaacc ttaccacaag aacgtttatca atctgctgga gcaagggtctt cctgtcctga 1740
60 tttacgcagg agacaaggat ttcatctgca attggctcgg aaaccaagcc tgggtccaatg 1800
61 agctcccttg gctctggacac gatgaattcg agtccccga gctgtacaac ctacacctga 1860
62 aggatggcac taaggtcggc gaggtcaaga atgctggcaa ttgtaccttt ctcaagaatgt 1920
63 ttgatggagg acacatgggt ccatacgacc agcctgagag ctctttggct atggtcaata 1980
64 gatggatagc ttgtgactac tccttgggaa ccaagaaata aaaagctcat gaaaacatta 2040
65 ttgcattggt gttcaactta tgtatgattg tttgttaaga tacactatct aaattctctg 2100
66 ttcaataacc agcctttcgg ttcttgtcc catttacata ttgatttgag catcgatcgc 2160
67 acaaccttga ctgttttttc gtttgagacg tcaatcttca gctcgtcaag aactgcag 2218
69 <210> SEQ ID NO: 2
70 <211> LENGTH: 2107
71 <212> TYPE: DNA
72 <213> ORGANISM: Hansenula polymorpha DL1
73 <300> PUBLICATION INFORMATION:
74 <308> DATABASE ACCESSION NO: U67173
75 <309> DATABASE ENTRY DATE: 1996-08-17
76 <313> RELEVANT RESIDUES: (1)..(2107)
77 <400> SEQUENCE: 2
78 gaggctcagct accttttggg agccatggca cgggtggcctt gattgtacgt tgctcctgca 60
79 atggtacaaa tgggattaca acaaaggtag cgtcaaatct cccgaaatag aagcaccgc 120
80 gatgagaaga gaaagtatat ctgcggagaa ttacatcaaa ctgaaaaagt caaagtcgcc 180
81 tgccacaggg tcagcatcca ttaaaatccc aatgtccccg catcacgttc gcaagctgag 240
82 tgaaaaacac ccaactgtcg ctatcgactt tcttctggac gactatatgg cgaaccacaa 300
83 ctttaacaag gcaaaaaaca gttcgtatgag gttttgaagc tgtcaaaacta cggatcgata 360
84 tcttaatcag atcaacttgc tattctggaa gccgcaccag tactagtggg gatagcgtaa 420
85 atgctcaaaa ctgatgtggt tatgtgtgac gcaactaccta aaatatgagg gtggcaaaaga 480
86 gagtataaatt tggggttcat taagttaggt ccataattta caaacctgtt ctgaatagag 540
87 agtcaccaat ccagttgcc caaaaccttg tcaaaatcat tagcatcgtg agcatctatt 600
88 ataaacaaag cgaatctcaa agatgaaact ctctttccca accctttact cgcttgcct 660
89 tgtgcttggg ttggtctccc tggccgatgc caaggtccac aaagctccca tcaaaaaagc 720
90 tctgcacag gctacttacc aggatgttac tgtcggcgat tatgttgagt cggtgaagca 780
91 aaagtatgtt acaacttaca acaagtttat cgccgcacca cagaatgacc agcaaatatt 840
92 ctcgatcgcc aagcgttctg acgagagtgc aagctcgggt cacaacactc ctttgaccaa 900
93 ctatctcaat gcccaatact tcaccgagat tcaattgggc actcctggcc aatcgttcaa 960
94 ggttatcctc gacaccggtt cctctaacct gtgggttctt agtagcgatt gcacgtcctt 1020

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/993,192

DATE: 07/30/2003
 TIME: 10:45:47

Input Set : N:\Crif3\RULE60\09993192.raw.txt
 Output Set: N:\CRF4\07302003\I993192.raw

```

95      ggccgtgttac ttgcatacta aatacgaacca cgatgagtcg tccacttacc agaagaacgg 1080
96      ttcttcgttt gctattcaat atggctcggg ttccctcgag ggttacgttt cccaagacac 1140
97      actgactatt ggtgaccttg tcatccctaa acaggatttc gccgaggcta ctacgagacc 1200
98      tggcttggct tttgcttttg gaaagtttga oggtattctg ggtctggcct acgacacgat 1260
99      ctcggtcaac agaatcgttc ctccaattta caatgctatc aatttgggat tgctggacac 1320
100     cccacaattc ggattttacc ttggtgacac ttccaagtcc gagcaggatg gaggagagggc 1380
101     tacctttggt ggatacgatg tgtctaagta cacaggcgat atcacctggg tgccagtcag 1440
102     aagaaaaggt tattgggaag tgaagtttag tggatatgcc ttggtgacga atacgctcca 1500
103     ttggagaaca ccgggagctg ccattgatac cggaaacctct ttgattgtct ttcactctca 1560
104     attggctgag attttgaact ctcaaatggg tgccgagaag tcatggctcg gacagtacca 1620
105     gatcgattgt gacaagagag actcaactgcc tgacctcact ttcaacttcg atggttacaa 1680
106     ctttaccatc tgcgcttatg actacacttt ggaagtttct ggttcttgta tttctgcatt 1740
107     cactccaatg gaacctccag ccccaattgg tccaatggcc atcattgggt acgctttcct 1800
108     cagaagatac tactctgtct atgaccttgg cagagacgct gtcggatttg ctaaggctgt 1860
109     gtaatgcaag gcagttacgt attagtttca acttgctagc atgaatttct atcgggtacg 1920
110     gttgcttgac aacgattttc tttgtgtctc tattactgct ataaaacatg aaaagattgg 1980
111     atttttttga cagtttctag taattgctta aatctgctct tttccttttg ttgatcccg 2040
112     gctgctctgt gtgcttcctt tttttatcta tattattaga ttgtctcttc tttagcagcg 2100
113     gaagctt                                     2107
115 <210> SEQ ID NO: 3
116 <211> LENGTH: 3550
117 <212> TYPE: DNA
118 <213> ORGANISM: Hansenula polymorpha DL1
119 <400> SEQUENCE: 3
120     gaattctata ggaatggtgc aaaacgaagg atcgtgtgga ctcatgaatg tgctatattt 60
121     gtggctaaaa agcctgccag agctcgcagg agataagagg tctggcgctg tgctgcgcaa 120
122     aatgatcgcc ttgataaaat tctcaatctc cgatcatgaa tcagggttca agtatttttg 180
123     tttttctttt caagacgagt atgaactgga acacgctctc gagtccetaa cgaccaggac 240
124     atcaagcagg cgccgcactc cgaagatctc gccgcttttg atttccaaca ctttgacgtc 300
125     gtggaggttg accagacctt ttctgatgcc tatgagctcg agctcggctc cgaacagggt 360
126     tcctgggtct aacgtgaatt tcaggttggt ggtaagagaa atcagcccca ctttgcgtga 420
127     ctcggttttg aaagcccgca ccagctgggt gaggcgtaaa agcgggacgg gatctgcctt 480
128     gttggcggtc tggatcgga ttggtggcac agattttctg tattgtttgt tgatcgatga 540
129     ttgaatgtac agtatgagcg ataacgagtc gttcgacttg ttgataaatt gtacgttcca 600
130     tttgaacacc tcgccgagct tgactttagt ctggcccga aactgaaatgc tgaggtttct 660
131     cttggagtgg ttgagagtca ccgacgagtt ggatcgtagt ttatttttct taggcaaat 720
132     cagagcagag ttggtcgact tggatttaaa taccocgaga ctaggcgacg cgattggggg 780
133     cacggtgggg gccgccataa aatccacatt ttagtccat cgtgtcaaga tttttctatt 840
134     ccggttgacc gtgcaatgga ttgtgacaaa aaccggttta acccaactga tgtcgttggt 900
135     tgttagtctg tacgtgagcg acagaataga atggttctgc aacaccaatg gatactttac 960
136     caccgagtag ccggcacaca gacagtttct gatgtccata tctaccgact cgattctcac 1020
137     cgacacttct gacgcccga cctgtccaga aagagccact tccaaagagg ccatggtttt 1080
138     ctccgtcttg atggtcttga accgcatctg gaatacgggg ctgatccgca gcgacgtgtc 1140
139     ggccgttacg agttcgtaaa ggtccacatg ctctgcagc acagtgttgc ctatagtac 1200
140     tgtatcctcg aactctaaat tttctatctt tctcgactgt tgcgaaacg aggtcttgat 1260
141     ctgcaattga ttgaaagt ttgtcaggcac cagggtcggt tccagtatcc atacggcctc 1320
142     agaccctatg acgacttttg aatcagcac ttttctcgct gtgttttgca agttgtatac 1380
143     caggtcgacc tcgactgtg tcacagacga cggagtgtac accacgactt ggatgtctc 1440
144     gtcgaatatg gcaatgtcca gcgtgtcgca ctggcgcagc tgttccacgg agcccagcac 1500

```

RAW SEQUENCE LISTING

DATE: 07/30/2003

PATENT APPLICATION: US/09/993,192

TIME: 10:45:47

Input Set : N:\Crif3\RULE60\09993192.raw.txt

Output Set: N:\CRF4\07302003\I993192.raw

```

145 atcagcagaa acccggtgtt tcggaacaag cacaccaaac attggaacaa acacttatcg 1560
146 acgcgggtgt gtttaaaaaa taatattaga tcgactgctc ctactctttt aaattcttta 1620
147 tttatccagc gcttagacag cacagagcta accatgagtc ggtatttttt actagcgtgc 1680
148 acattggtct tgcaatgtgt cgcggcctcc caggaggact acatagtcaa ggatttgctt 1740
149 ggactctcga atattcctgc cgtggtgagg ccagtgatgc atgccggaca ccttgaaata 1800
150 gatgaggaac acaacaccga gctgtttttc tggcgattcc agaatccgaa gaacaacggg 1860
151 acacacaaa cgtccaccg caatgagctg atcgtctggt tgaacggtgg ccttggtctg 1920
152 tcgtcgatgg atggagccat gatggaaaca ggccccctgc ggtgtgcaga caagttggag 1980
153 gtggagctca acccggtatc ctggacacag gtggccgata tcttgtttgt ggatcagccg 2040
154 gccggaattt tgagaaagtg gagtgcgtca aaatatcca tagtattttg gctgccagta 2100
155 gagagaaac caagcctgcc aaagaacaat gtgtcaacat gtacgactac cgcaaacatg 2160
156 attacttccc tgcattgcca tccaattggc cagaaggcct gccaccgtg actaaatttc 2220
157 tgaatctgga tgcgtgcaa aaagccctga acctgaaatc ggcgaagaga tggcacgagt 2280
158 gtgacggaaa agtcgaattc tttttccagc cagagcactc tgtcaagtcg ttcgacctgc 2340
159 tgccaaaact actggagaaa atgaaaatcg cgctgtttgc ttggcgacaag gacattatct 2400
160 gcaaccacaa atcaatagaa atggtgattg aaaaattgca aattacacca ggacaattcg 2460
161 gtttcacaaa cagcagaaag tctggctgga tctacgacgg gcaggaagtg ggtgaggttg 2520
162 agacacgtcg gaacctcacg tatatcaagg tgttcaattc gtgcataatg gtgccttacg 2580
163 acctgccgga ggtgagcaga ggactattcg atataatcac aaactctatt gaaaaacggt 2640
164 cgacagacat tgtgacgccc gtttatgaca gcagaggcaa ctacaagttt gtggaggaga 2700
165 agcaggacac cgaccagaac gaggaagaag agaaggagaa acctcccaag caccatcata 2760
166 gtctcacgtt ttacgtggca gaggtggcga ttctggcagt gctagcatat ctgctttaca 2820
167 gcttctacaa atcgttcgcc aaatcgcgta agtctgcatt tttgtcactc tcttccaaga 2880
168 agaagaagaa gcaggtgcac tggtttgacg agagcgacat aggaatggac caggaagctg 2940
169 gcgaggcaga tcataagcct aagagcatgc tggagtcggt gttcaacaag ctgggctatg 3000
170 gaggacagta tgacacggtg caggacggcc gtgacatcga gatggcgccg gtcgaagaac 3060
171 acgaagacca atttattatt caaagcgacg aggaagagtt tggccacaga tagagtacga 3120
172 ttatataata tccacgaaaa attagtttcc aattatttcg ctccctgattg taatgtctca 3180
173 aagttggatc ctctggaagc ggtccctgat taccggcggt ggaattattg gaagcggtgt 3240
174 gctactgtac aaattcacga ctctaccga ggaacagttg atcgctaaat tatccccga 3300
175 gctgagagca gactatgaac gcaacaaga attgcgaggg aaagaacagg agatgttgat 3360
176 ggagatcgtc aaacagacgg cagcgtcgaa cgatccggtg tggaaagacg ggccgattgt 3420
177 gtcgcggtgg gaccgggact ttacgccatc gagggaaagt ttattagtga agagagagcg 3480
178 atttgaaaaa gagcaggcgg agaaaaaaca gcgcgaagag ctcgagcgtc tgaagccga 3540
179 ggccaagctt                                     3550

```

```

181 <210> SEQ ID NO: 4
182 <211> LENGTH: 17
183 <212> TYPE: DNA
184 <213> ORGANISM: Artificial Sequence
185 <220> FEATURE:
186 <223> OTHER INFORMATION: PCR primer for S. cerevisiae PRC1 gene
187 <400> SEQUENCE: 4
188 atgaaagcat tcaccag
190 <210> SEQ ID NO: 5
191 <211> LENGTH: 17
192 <212> TYPE: DNA
193 <213> ORGANISM: Artificial Sequence
194 <220> FEATURE:
195 <223> OTHER INFORMATION: PCR primer for S. cerevisiae PRC1 gene

```

17

RAW SEQUENCE LISTING DATE: 07/30/2003
 PATENT APPLICATION: US/09/993,192 TIME: 10:45:47

Input Set : N:\Crf3\RULE60\09993192.raw.txt
 Output Set: N:\CRF4\07302003\I993192.raw

```

196 <400> SEQUENCE: 5
197     ttataaggag aaaccac                                     17
199 <210> SEQ ID NO: 6
200 <211> LENGTH: 17
201 <212> TYPE: DNA
202 <213> ORGANISM: Artificial Sequence
203 <220> FEATURE:
204 <223> OTHER INFORMATION: PCR Primer for S. cerevisiae PEP4
205 <400> SEQUENCE: 6
206     atgttcagct tgaaagc                                     17
208 <210> SEQ ID NO: 7
209 <211> LENGTH: 16
210 <212> TYPE: DNA
211 <213> ORGANISM: Artificial Sequence
212 <220> FEATURE:
213 <223> OTHER INFORMATION: PCR primer for S. cerevisiae
214 <400> SEQUENCE: 7
215     tcaaattcgt ttggcc                                     16
217 <210> SEQ ID NO: 8
218 <211> LENGTH: 27
219 <212> TYPE: DNA
220 <213> ORGANISM: Artificial Sequence
221 <220> FEATURE:
222 <223> OTHER INFORMATION: PCR primer for H. polymorpha KEX1
223 <400> SEQUENCE: 8
224     tgggysaacg ghccwgggtg ytcbtcb                         27
226 <210> SEQ ID NO: 9
227 <211> LENGTH: 30
228 <212> TYPE: DNA
229 <213> ORGANISM: Artificial Sequence
230 <220> FEATURE:
231 <223> OTHER INFORMATION: PCR primer for H. polymorpha KEX1
232 <400> SEQUENCE: 9
233     wggratgtay tgwccrgcgt avgactcdcc                       30

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/30/2003
PATENT APPLICATION: US/09/993,192 TIME: 10:45:48

Input Set : N:\Crf3\RULE60\09993192.raw.txt
Output Set: N:\CRF4\07302003\I993192.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 10

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/993,192

DATE: 07/30/2003

TIME: 10:45:48

Input Set : N:\Crf3\RULE60\09993192.raw.txt

Output Set: N:\CRF4\07302003\I993192.raw